Cross-Layer Multi-User Selection in 5G Heterogeneous Networks Based on Hybrid Beamforming Optimization for Millimeter-Waves

Ahmad Fadel, Ahmad Nimr, Hsiao Lan Chiang, Marwa Chafii, Bernard Cousin

Cross-Layer Multi-User Selection in 5G Heterogeneous Networks Based on Hybrid Beamforming Optimization for Millimeter-Wave

Ahmad Fadel[1], *Ahmad Nimr*[2], *Hsiao – Lan Chiang*[2], *Marwa Chafii*[3], *Bernard Cousin*[1]
[1] IRISA University of Rennes, France
{ahmad.fadel, bernard.cousin }@irisa.fr
[2] Vodafone Chair Mobile Communication Systems, Technische Universität Dresden, Germany
{ahmad.nimr, hsiao-lan.chiang}@ifn.et.tu-dresden.de
[3] ETIS UMR 8051, Université Paris-Seine, Université Cergy-Pontoise, ENSEA, CNRS, France
marwa.chafii@ensea.fr

***Abstract*—Lack of coordination between network layers limits the performance of most proposed solution for new challenges posed by wireless networks. To overcome such limitations, cross-layer physical and medium access (PHY-MAC) design for multi-input-multi-output orthogonal frequency division multiple access system in heterogeneous networks (HetNETs) is proposed. In this paper, we formulate an optimization problem for hybrid beamforming, in a multi-user HetNET scenario aiming to maximize the total system throughput. Furthermore, analog beamforming is selected from a codebook containing a limited number of candidates for steering vectors. The proposed problem is non-convex and hard to solve. Thus it is relaxed by transforming it into a subtraction form of two convex funcions. Afterward we apply a group of well-known metaheuristic algorithms to calculate the normalized hybrid beamforming vectors. The optimal solution is obtained using an exhaustive search (ES) algorithm that provides an ideal solution, but with high complexity. In addition, zero-forcing-based approach (ZFA), matched filter (MF), and QR-based approach (QR) are applied to get quick sub-optimal solutions. Hence, we analyze the performance of our systems using the throughput metric. The simulation results show that QR algorithm outperforms ZFA and MF in low and middle signal-to-noise ratio (SNR) regime, while ZFA outperforms QR and MF at higher SNRs. Moreover, QR is close to the optimal solution ES.**

***Index Terms*—Cross-layer, user selection, 5G heterogeneous network, beamforming, millimeter waves, orthogonal steering vector.**

I. Introduction

One of the main 5th generation (5G) requirements is to support 1000 times larger capacity per area compared with current Long Term Evolution (LTE) technology, but with similar cost and energy dissipation per area as today's cellular systems. In addition, an increase in capacity will be possible if all the three factors that jointly contribute to system capacity are increased; 1) more spectrum using millimeter waves (mmWaves) spectrum band, 2) a large number of base stations per area by mean of densification, and 3) an increased spectral efficiency per cell [1]. Massive multi-input-multi-output (MIMO) systems are considered essential in contributing to the latter factor, as they promise to provide a highly increased spectral efficiency per cell. Indeed, it takes advantage of the spatial degrees of freedom (DoF) that provide spatial multiplexing that inherently inherently minimize intra-cell and inter-cell interferences [2]. Obviously, this can be achieved by applying the hybrid beamforming technique knowing that the system operates at mmWaves frequency bands. Thus, the gain realized through antenna beamforming can compensate for the high path at mmWaves frequencies. Accordingly, a combination of analog beamforming (operating in passband) and digital beamforming (operating in baseband) can be one of the low-cost solutions [3]. This is because using only digital beamforming requires more radio frequency (RF) chains, which leads to high implementation cost and power consumption. An effective beam finding mechanism is important in mmWaves communications. Accordingly, generating a codebook that is composed of limited steering angles and steering vectors could incarnate numerically the physical beams. It was reported in [4] that using orthogonal steering vectors provide higher data rates, because the spatial frequencies of the given angle-of-arrivals (AoAs) and the orthogonal steering angles have nearly the same distribution.

In a traditional network, the optimization is usually carried out considering a respective layer objectives based on only local information ignoring other layers' design parameters or information. This fact gives a locally optimal, but globally suboptimal solution. Cross-layer design refers to sharing information among layers for efficient use of network resources and achieving high adaptivity. This motivates us to formulate an optimization problem for a cross-layer design, having the hybrid beamforming as an optimization variable, knowing that the physical layer (PHY) is responsible for signaling and channel estimation, whereas medium access layer (MAC) is responsible for resource allocation and multi-user selection.

In the literature, the authors in [5] focused on optimal analog and digital beamforming designs in a multi-user beamforming scenario to study the impact of energy efficiency on spectrum efficiency, and they showed that hybrid beamforming achieves better channel estimation performance than the method solely based on analog beamforming. But they address only PHY layer to avoid complexity when dealing with a cross-layer design. The authors in [6], [7], formulated an optimization problem to determine the hybrid beamforming in the downlink (DL) scenario, having a backhaul and a power constraint. Their formulation actually aims to maximize the throughput. But, the optimal solution was the standard zero-forcing technique without having any normalization which leads to an increase in power and resulting in a non-accurate outcome. In our previous work [3], we have only investigated PHY hybrid beamforming optimization which has been based on an implicit channel state of information for mmWaves links. Zero forcing ZFS, and QR-decomposition algorithms were applied in a single cell multi-user selection while using statistical channel model where path loss is normalized and the study is based on the 3rd technology of wireless communication networks.

In this paper, we address the issue of cross-layer in a heterogeneous network, where the PHY and MAC knowledge of the wireless medium is shared, to provide a hybrid beamforming optimization for a multi-user scenario. In order to meet the overwhelming demands of network throughput for a practically important case wherein the number of users, N_u is larger than the number of transmit antennas N_t, as we propose to select Q users among N_u to attribute for them the resources. Our contributions are six-folds: i) Generate a new system model to support cross-layer design for multi-user selection, markedly it consists of both analog and digital beamforming at the transmission side. ii) Formulate an optimization problem for a cross-layer design, having the hybrid beamforming as an optimization variable. Indeed, the heterogeneous cellular network is based on two different technologies (4th generation (4G) and 5G), and the purpose of our formulation is maximizing the system throughput; iii) Transform the optimization problem in order to relax the non-convexity by applying

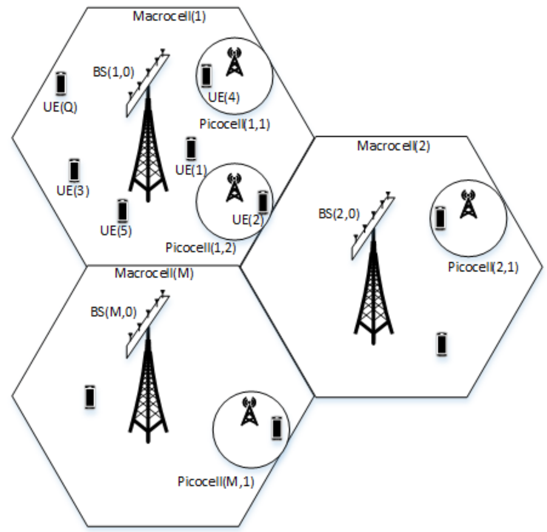

Figure 1. Heterogeneous network with MIMO base stations.

the difference-of-convex functions (DC) programming [8]; iv) Compute the optimal solution by applying the ES algorithm that will be considered as the ideal solution, and has been proposed in our previous work [9]; v) Propose a sub-optimal solution to reduce the complexity and produce solutions close to the optimal one; vi) Assess the performance of zero-forcing-based approach (ZFA), QR-based approach (QR) [7] and matched filter (MF) approach compared to that of exhaustive search (ES), versus throughput evaluation metric.

The rest of the paper is organized as follows: Section II outlines the proposed system and channel models. In Section III we formulate hybrid beamforming with a fixed analog beamforming optimization. Section IV describes the optimal and reduced-complexity algorithms. Furthermore, Section V provides a simulation-based comparison of the system throughput performance for ES, ZFA, QR and MF algorithms with respect to SNR and number of users. Conclusions and future work are mentioned in Section VI.

II. HetNet Multi User MIMO-OFDMA Network Model

A. Heterogeneous Network Model

In a heterogeneous multi-user MIMO-OFDMA network, we assume to have M macrocells overlaid by P picocells with a N_u user equipment (UEs) distributed overall the system area as depicted in Fig. 1. Indeed $Q \leq N_u$ (UEs) will be selected to be served. Each base station (BS) is equipped with N_t transmit antennas, and each UE with $N_r = 1$ receive antenna.

One of the most attracted characteristics in MIMO systems is the spatial multiplexing, literally because each BS can serve up to $\frac{N_t}{N_r}$ UEs simultaneously for each radio ressource unit. The intra-cell interference is

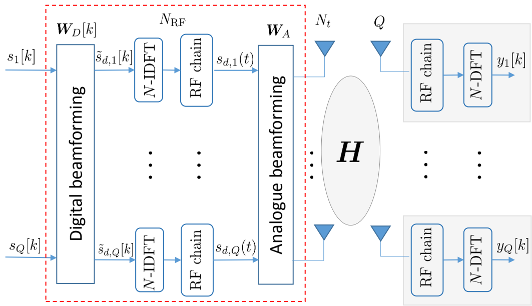

Figure 2. Multi-user System Model with Analogue and Digital Beamforming.

roughly negligible, due to the fact of associating the 4G frequency spectrum to the macrocell, while associating the mmWaves that fall in the spectrum range of 5G to picocells. Moreover, picocells are far enough from each other in order not to affect each other by any kind of interference. Thanks to fiber optic backhaul links, macro and pico BSs are connected to a centralized control unit. We assume that the wireless channel operates in time-division-multiplexing (TDD) system that relies on reciprocity, by which the uplink channel is used as an estimate of the downlink channel, and this occurs when receiving a pilot training sequence from terminal devices toward the base station where channel state information (CSI) is obtained.

B. MULTI USER MIMO-OFDMA System Model

In this subsection we consider only one BS with N_{t} transmit antennas employing orthogonal frequency division multiplexing (OFDM) system. By means of digital beamforming (DBF) and analog beamforming (ABF), $Q \leq N_{\mathrm{t}}$ users can be simultaneously served with the same time and frequency resources. Following the multiuser (MU)-MIMO scheme as illustrated in Fig. 2, let $\boldsymbol{s}[k] = [s_1[k], \cdots, s_Q[k]]^T \in \mathbb{C}^{Q\times 1}$ be the data symbol vector transmitted on the k-th subcarrier. The user data are precoded with a DBF matrix $\boldsymbol{W}_D[k] \in \mathbb{C}^{Q\times N_{\mathrm{RF}}}$, such that

$$\boldsymbol{s}_d[k] = \boldsymbol{W}_D[k]\boldsymbol{s}[k] \in \mathbb{C}^{N_{\mathrm{RF}}\times 1} \tag{1}$$

which is a precoded data vector. The precoded data are transformed to the time domain with inverse discrete Fourier transform (IDFT) transform and passed to N_{RF} RF chains, to generate the analog signal vector $\boldsymbol{s}_d(t) \in \mathbb{C}^{N_{\mathrm{RF}}\times 1}$. After ABF with a matrix $\boldsymbol{W}_A \in \mathbb{C}^{N_{\mathrm{RF}}\times N_{\mathrm{t}}}$, the transmitted signal vector is given by

$$\boldsymbol{x}(t) = \boldsymbol{W}_A\boldsymbol{s}_a(t) \in \mathbb{C}^{N_{\mathrm{t}}\times 1}. \tag{2}$$

assuming that a cyclic prefix (CP) of sufficient length is inserted, and let $\boldsymbol{H}(t) \in \mathbb{C}^{N_{\mathrm{t}}\times Q}$ be the MU-MIMO channel states between each pair of N_{t} transmitter antennas and each Q users. The received signal $y_q(t)$ of the q-th user can be expressed with the circular convolution

$$y_q(t) = \sum_{n_t=1}^{N_{\mathrm{t}}} [\boldsymbol{H}(t)]_{(q,n_t)} \circledast [\boldsymbol{x}(t)]_{(n_t)} + v_q(t), \tag{3}$$

where $v_q(t)$ is the additive white Gaussian noise (AWGN). After the sampling of $y_q(t)$ and performing N-discrete Fourier transform (DFT), the circular convolution is translated to element-wise product with the channel coefficients in the frequency domain, which are defined by the channel matrix $\tilde{\boldsymbol{H}}[k] \in \mathbb{C}^{N_{\mathrm{t}}\times Q}$. Thus, we get the signal

$$y_q[n] = \sum_{n_t=1}^{N_{\mathrm{t}}} \left[\tilde{\boldsymbol{H}}[k]\right]_{(q,n_t)} [\tilde{\boldsymbol{x}}[k]]_{(n_t)} + \tilde{v}_q[k], \tag{4}$$

where $\tilde{\boldsymbol{x}}[k] = \boldsymbol{W}_A\boldsymbol{W}_D[k]\boldsymbol{s}[k]$ is the consequence of (1) and (2). The effective channel that can be regarded as a coupling of the channel having analog beamforming gain on both sides is defined by

$$\boldsymbol{H}^{(e)} = \tilde{\boldsymbol{H}}\boldsymbol{W}_A \in \mathbb{C}^{Q\times N_{\mathrm{RF}}}, \tag{5}$$

assuming that the ABF is fixed during at least one OFDM symbol. The corresponding multiple-input single-output (MISO) channel of the q-th user is defined by

$$\boldsymbol{h}_q^H[k] = \left[\boldsymbol{H}^{(e)}\right]_{(q,:)} \in \mathbb{C}^{1\times N_{\mathrm{t}}} \tag{6}$$

$$\text{then, } y_q[k] = \boldsymbol{h}_q^H[n]\boldsymbol{W}_D[k]\boldsymbol{s}[k] + \tilde{v}_q[k]. \tag{7}$$

The goal is to find the DBF matrix that maximizes the sum rate for the k-th subcarrier. First, we define the normalized DBF matrix $\bar{\boldsymbol{W}}[k] = [\boldsymbol{w}_1, \cdots \boldsymbol{w}_Q] \in \mathbb{C}^{N_{\mathrm{t}}\times Q}$, $\|\boldsymbol{w}\|^2 = 1$ as

$$\boldsymbol{W}_D[k] = \bar{\boldsymbol{W}}[k]\boldsymbol{\Lambda}[k], \tag{8}$$

with $\boldsymbol{\Lambda}[k] \in \mathbb{C}^{Q\times Q}$ is a diagonal matrix that defines the power allocation such that

$$[\boldsymbol{\Lambda}[k]]_{(q,q)} = \| [\boldsymbol{W}_D[k]]_{(:,q)} \| = \sqrt{P_q[k]}. \tag{9}$$

From now on, the subcarrier index k is dropped for the simplicity of notation. In order to compute $\boldsymbol{W}_D$, first we need to find the unit-norm vectors $\{\boldsymbol{w}_q\}$ in addition to P_q. We assume that the data samples are uncorrelated with power E_s, and the power constraint $\sum_{q=1}^{Q} P_q = Q$ must be fulfilled. The received signal model can be rewritten as

$$y_q = \sqrt{P_q}\boldsymbol{h}_q^H\boldsymbol{w}_q s_q + \underbrace{\sum_{i\neq q} \sqrt{P_i}\boldsymbol{h}_q^H\boldsymbol{w}_i s_i}_{\text{IUI}} + \tilde{v}_q. \tag{10}$$

C. Channel Model

We considered a heterogeneous network, and the proposed system model is applied for each macro and picocell. Let $\boldsymbol{h}_{nq} \in \mathbb{C}^{1\times N_{\mathrm{t}}}$ be the channel vector between the (m,p)-th BS and the q-th user equipment (UE) of the (m',p')-th cell in the k-th resource unit.Here the shorthand notation $n=(m,p)$ is used for simplicity. Where m and p are the indices of the distributed cells in the network, knowing that $m \in \{1,\dots,M\}$ and $p \in \{0,\dots,P\}$, for clarification $p=0$ corresponds to the BS macrocell. Then, the received signal at the q-th UE of the (m',p')-th cell in the k-th resource unit, y_{nq}, is given by

$$\begin{aligned}y_{nq}(k)=&\sqrt{P_{nq}(k)}\boldsymbol{h}(k)_{nq}^{H}(k)\boldsymbol{w}_{nq}(k)s_{nq}(k)\\ &+\underbrace{\sum_{i\neq q}\sqrt{P_{ni}(k)}\boldsymbol{h}_{nq}^{H}(k)\boldsymbol{w}_{ni}(k)s_{ni}(k)}_{\text{IUI}}+\tilde{v}_{nq}(k),\\ &\forall\ n=(m,p)\,with\,m\in[1\dots M]\,and\,p\in[0\dots P]\end{aligned} \tag{11}$$

Moreover, $\boldsymbol{h}_{nq}(k) = l_{nq}(k)g_{nq}(k)\tilde{\boldsymbol{h}}_{nq}(k)$, where $\boldsymbol{l}_{nq}(k)$, $\boldsymbol{g}_{nq}(k)$, denotes the path loss and shadowing, respectively. The channel is modelized using Rapapport model [10], which take into account path loss and shadowing. $\tilde{\boldsymbol{h}}_{nq}[k]$ is the small scale fading coefficients, which can be generated using the statistical channel model [11]. In addition, $\tilde{v}_{nq}\sim\mathcal{N}(0,\sigma_{nq}^2)$ is the AWGN noise and $I_{nq}(k)$ is inter-user interference (IUI).

III. Problem Formulation

We formulate an optimization problem for a heterogeneous network, aiming to optimize hybrid beamforming techniques. We have used the mmWaves propagation characteristic for picocells to maximize the total average system throughput, due to the large frequency spectrum range.

A. System Performance Evaluation Metric

We consider the system throughput as the performance metric in this paper. The instantaneous channel throughput $R_{nq}(k)$ for the q-th user of the n-th cell in the k-th resource unit is given by

$$\begin{aligned}&R_{nq}(k)=B_w\log_2(1+SINR_{nq}),\\ &\forall n=(m,p)\,with\,m\in[1\dots M]\,and\,p\in[0\dots P]\end{aligned} \tag{12}$$

where B_w is the channel bandwidth, and SINR_{nq} denotes the signal-to-interference-plus-noise ratio (SINR) for the q-th user, which is given by:

$$\text{SINR}_{nq}=\frac{E_sP_{nq}\boldsymbol{w}_{nq}^H\boldsymbol{h}_{nq}\boldsymbol{h}_{nq}^H\boldsymbol{w}_{nq}}{\sum\limits_{i\neq q}\left[E_sP_i\boldsymbol{w}_{ni}^H\boldsymbol{h}_{nq}\boldsymbol{h}_{nq}^H\boldsymbol{w}_{ni}\right]+\sigma_{nq}^2}, \tag{13}$$

The average system throughput, $\mathbb{V}$ in bit/s/Hz/BS is defined by

$$\mathbb{V}(\{\boldsymbol{w}_{nq}\},\{P_{nq}\})=\frac{1}{M*(P+1)}\sum_{m=1}^{M}\sum_{p=0}^{P}\sum_{q\in Q_n}\sum_{k=1}^{K}R_{nq}(k) \tag{14}$$

where $\{\boldsymbol{w}_{nq}\}$ and $\{P_{nq}\}$ refer to the set of unit-norm vector of the digital beamforming matrix and power allocation of the served users, respectively. The number of served users in the n-th macro and pico BS is denoted as Q_n $\forall n=(m,p)$.

B. DBF Optimization Problem

After Q_n the users in the n-th cell are selected, the DBF needs to be optimized. Because the logarithm function in (12) is an increasing function, then maximizing $R_{nq},\ q=1,\cdots,Q_n$ is equivalent to maximizing SINR_{qn}. Therefore, the DBF optimization problem can be written as:

$$\begin{aligned}&\max\ \text{SINR}_q\left(\{\boldsymbol{w}_q,P_q\}\right)\\ &\text{s.t. } \|\boldsymbol{w}_q\|=1,\ \sum_{q=1}^{Q}P_q\leq Q,\ P_q>0,\ q=1,\cdots,Q.\end{aligned} \tag{15}$$

Note that, the index n is dropped for simplicity. For a given $\{\boldsymbol{w}_q\}$ the problem turns to finding $\{P_q\}$, i.e. solving a power allocation problem.

C. Power Allocation Problem

Let $\Omega_{q,i} = \boldsymbol{w}_i^H\boldsymbol{h}_q^H\boldsymbol{h}_q\boldsymbol{w}_i$, and $\boldsymbol{p}\in\Re^{Q\times 1}$, where $[\boldsymbol{p}]_{(q)}=P_q$. Then

$$\begin{aligned}R(\boldsymbol{p})&=\sum_{q=1}^{Q}\log_2\left(1+\frac{E_sP_q\Omega_{q,q}}{\sum\limits_{i\neq q}[E_sP_i\Omega_{q,i}]+\sigma_q^2}\right)\\ &=\sum_{q=1}^{Q}\log_2\left(\frac{1+\frac{E_s}{\sigma^2}\sum\limits_{q=1}^{Q}P_i\Omega_{q,i}}{1+\frac{E_s}{\sigma^2}\sum\limits_{i\neq q}P_i\Omega_{q,i}}\right)=\sum_{q=1}^{Q}\log_2\left(\frac{1+\boldsymbol{a}_q^T\boldsymbol{p}}{1+\boldsymbol{b}_q^T\boldsymbol{p}}\right)\end{aligned} \tag{16}$$

$$\begin{aligned}\text{where },\ &[\boldsymbol{a}_q]_{(i)}=\Omega_{q,i},\ i=1\cdots Q\\ &[\boldsymbol{b}_q]_{(i)}=\Omega_{q,i},\ i=1\cdots Q,\ i\neq q.\end{aligned} \tag{17}$$

Thus, the sum rate maximization can be written as

$$\begin{aligned}&\max_{\boldsymbol{p}}\ R\\ &\text{S.t. } R\leq\sum_{q=1}^{Q}\log_2\left(\frac{1+\boldsymbol{a}_q^T\boldsymbol{p}}{1+\boldsymbol{b}_q^T\boldsymbol{p}}\right),\ P_q>0,\ \sum_{q=1}^{Q}P_q=Q.\end{aligned} \tag{18}$$

This problem is a difference between two convex functions. First, the problem is reformulated as

$$\begin{aligned}&\max_{\boldsymbol{p}}\ R,\\ &\text{S.t. } R\leq\sum_{q=1}^{Q}\log_2\left(1+\boldsymbol{a}_q^T\boldsymbol{p}\right)-t,\\ &\sum_{q=1}^{Q}\log_2\left(1+\boldsymbol{b}_q^T\boldsymbol{p}\right)\leq t,\ P_q>0,\ \sum_{q=1}^{Q}P_q\leq Q,\end{aligned} \tag{19}$$

Using the linear approximation to convert the second constraint, then we get the convex problem

$$\begin{aligned}
&\max_{\boldsymbol{p}} R,\\
&\text{S.t.}\ \ R \leq \sum_{q=1}^{Q} \log_2\left(1+\boldsymbol{a}_q^T\boldsymbol{p}\right) - t,\\
&\qquad \sum_{q=1}^{Q}\left(\log_2\left(1+\boldsymbol{b}_q^T\boldsymbol{p}_0\right) + \frac{\boldsymbol{b}_q^T}{1+\boldsymbol{b}_q^T\boldsymbol{p}_0}\left[\boldsymbol{p}-\boldsymbol{p}_0\right]\right) \leq t,\\
&\qquad P_q > 0,\ \sum_{q=1}^{Q} P_q \leq Q.
\end{aligned} \tag{20}$$

This problem can be solved iteratively; first we set the initial power allocation $\boldsymbol{p}_0$ to a uniform allocation, and this vector is updated after each iteration. The algorithm stops when $|R_{i+1} - R_i| < R_i\epsilon$, where ϵ defines the change threshold.

IV. Optimal and Low-Complexity Algorithms

In this section, we present the analog beamforming selection phase, and we apply the four algorithms ES, ZFA, MF, and QR that take into consideration the same constraints and objective of the optimization problem. Accordingly, the distributed power between users should be less or equal the maximum power in each BS. The number of users Q that share the same resource unit k should not exceed the number of transmitted antenna N_t installed on a BS.

A. Analog Beamforming Selection

A codebook based beamforming training procedure can balance the trade off between complexity and high performance. The N_{RF} analog beamforming vectors of $\boldsymbol{W}_A$ in Fig. 2 are selected from a predefined orthogonal codebook $\mathcal{F} = \{\tilde{f}_{n_f} \in \mathbb{C}^{N_t\times 1}, n_f = 1,\ldots,N_F\}$ with the n_f^{th} member given by [12]

$$\tilde{f_{n_f}} = \frac{1}{\sqrt{N_t}}[1, e^{\frac{j2\pi}{\lambda_0}\sin(\phi_{n_f})\triangle_d},\ \ldots\ , e^{\frac{j2\pi}{\lambda_0}\sin(\phi_{n_f})(N_t-1)\triangle_d}]^T \tag{21}$$

where ϕ_{n_f} stands for the n_f^{th} candidate of the steering angles at the transmitter, $\triangle_d = \frac{\lambda_0}{2}$ is the distance between two neighboring antennas, and λ_0 refers to the wavelength for a specific carrier frequency.

1) Initial analog beam selection: This step is achieved by transmitting known pilot signals, and at the receiver they will include the effect of analog beamforming, then an observation used for the analog beam selection at subcarrier k for a specific beam $\tilde{f}_{n_f}$ can be acquired by correlating the k^{th} received pilot with its transmitted signal [3] as shown in the following equation

$$y_{n_f}[k] = \tilde{f}_{n_f}\boldsymbol{H}[\boldsymbol{k}]^{(e)} + AWGN \tag{22}$$

The idea from using the beamforming technique is to achieve the maximal signal-to-noise ratio (SNR) [13], and due to the assumption that the noise is signal-independent, the steering vector index can be selected individually and sequentially according to the sorted received energy estimates

$$\tilde{f} = \underset{\tilde{f}_{n_f}\in\mathcal{F}\backslash\mathcal{F}'}{\arg\max} \sum_{k=0}^{K-1} \mid y_{n_f}[k] \mid^2 \tag{23}$$

having $\mathcal{F}'$, is a cumulative set where the selected steering vectors are stored.

2) Steering angles selection: Having a limited codebook size, the way of designing the steering angles has a consequence effect on the beamforming performance, to compensate for the angles of arrival and departure (AoAs/AoDs). Therefore the steering angles are selected uniformly between the range $(-\frac{\pi}{2}, \frac{\pi}{2})$.

B. Exhausitve Search Algorithm using Standard Zero-forcing

One of the most important targets of telecommunication operators in the next generation of cellular networks is to maximize the system throughput. To put it another way, higher throughput means replying quickly to user demands, thus it achieves their satisfaction.

The objective of the optimization problem is to optimize a hybrid beamforming for a selected $Q \leq N_t$ users and respecting all the aforementioned constraints in a way to maximize the system throughput, while using hybrid beamforming (fixed analog and ZF for digital). It is important to realize that, getting the optimal solution by applying the exhaustive search algorithm, would be ideal to achieve the goal. Precisely, this algorithm checks all the possible combinations in a way to get the optimal solution by selecting a set of users that achieve the maximum total throughput, but it may have a severe drawback, the computational cost of ES may introduce a very long delay when the combinatorics of the problem is high.

In Algorithm 1, we propose to select $q \in Q$ from N_u users, by trying all the possible combinations to achieve the maximum total throughput. Power is uniformly distributed among the selected users.

C. Normalized beamforming vectors

Getting this optimal solution is of high complexity. Thus, computing normalized beamforming vectors to solve the hybrid beamforming optimization problem sounds a good solution. Accordingly, three cases of $\{\boldsymbol{w}_q\}$ are to be compared.

1) Zero-Forcing based approach: In this approach, the interference is canceled such that

$$\boldsymbol{H}^{(e)}\boldsymbol{W}_D = \boldsymbol{I}_Q \text{ thus}, \boldsymbol{W}_D = \boldsymbol{H}^{(e)H}\left[\boldsymbol{H}^{(e)}\boldsymbol{H}^{(e)H}\right]^{-1}. \tag{24}$$

The normalized vector is given by

$$\boldsymbol{w}_q = \frac{[\boldsymbol{W}_D]_{(:,q)}}{\|\,[\boldsymbol{W}_D]_{(:,q)}\,\|},\ \text{SINR}_q = \frac{E_s P_q \boldsymbol{w}_q^H\boldsymbol{h}_q\boldsymbol{h}_q^H\boldsymbol{w}_q}{\sigma_q^2}. \tag{25}$$

Algorithm 1 Exhaustive search [9]

Input: N_u, N_t, P_{max}

Output: The set $\mathbf{S_r}$

while $Q \in C_{N_u}^{N_t}$ **do**

  Compute $\forall q \in Q,\ p_q = \frac{P_{max}}{N_u}$

  Compute $\forall q \in Q,\ \boldsymbol{H}^{(e)} = \boldsymbol{H}\boldsymbol{W}_A$ Effective channel

  $\forall q \in Q,\ \boldsymbol{w}_q = H^*(HH^*)^{-1}$ ZF-beamforming

  $\forall q \in Q,\ SINR(q)$ (13)

  $R(c) = \sum_{\forall q \in Q} R(q)$ (12)

  if $R(S_r) < R(Q)$ **then**

    Let $S_r = Q$

  end if

end while

As a result, the sum rate optimization problem is reduced to water filling. The normalization preserves the power constraint, unlike the standard zero-forcing (ZF) approach, which may lead to the increase of the total power.

2) Matched filter: In this approach, the nominator of SNR_q is maximized, such that $\boldsymbol{w}_q = \frac{\boldsymbol{h}_q}{\|\boldsymbol{h}_q\|}$ However, the interference can be sever, and it is simple in implementation knowing that it is useful for low SNR regime.

3) Compromised interference QR approach: In this approach, first we sort the users according to the maximum $\boldsymbol{h}_q^H \boldsymbol{h}_q$. Then, we compute $\boldsymbol{w}_q$ as follows: the solution for $q = 1$ is given by $\boldsymbol{w}_1 = \frac{\boldsymbol{h}_1}{\|\boldsymbol{h}_1\|}$. The solution for $q > 1$ is achieved by solving

$$\begin{aligned} &\max\ \boldsymbol{w}_q^H \boldsymbol{h}_q \boldsymbol{h}_q^H \boldsymbol{w}_q \\ &\text{S.t. } \|\boldsymbol{w}_q\| = 1,\ \boldsymbol{h}_1^H \boldsymbol{w}_q = 0, \cdots, \boldsymbol{h}_{q-1}^H \boldsymbol{w}_q = 0. \end{aligned} \tag{26}$$

Therefore, we reduce the interference from the later users. Actually, the optimal solution can be written as $\boldsymbol{H}^{(e)}\bar{\boldsymbol{W}} = \boldsymbol{T}$, where $\boldsymbol{T} \in \mathbb{C}^{Q\times Q}$ is lower triangular matrix. By computing the QR decomposition of $\boldsymbol{H}^{(e)H}$ under the assumption that $N_\text{t} \geq Q$ such that $maH^{(e)H} = \boldsymbol{Q}\boldsymbol{R} \in \mathbb{C}^{N_\text{t}\times Q}$, where $\boldsymbol{R} \in \mathbb{C}^{Q\times Q}$ is an upper triangular matrix, and $\boldsymbol{Q} \in \mathbb{C}^{N_\text{t}\times Q}$ is orthogonal matrix, $\boldsymbol{Q}^H\boldsymbol{Q} = \boldsymbol{I}_Q$, thus,

$$\boldsymbol{H}^{(e)}\bar{\boldsymbol{W}} = \boldsymbol{R}^H\boldsymbol{Q}^H\bar{\boldsymbol{W}} = \boldsymbol{T}. \tag{27}$$

By choosing $\bar{\boldsymbol{W}} = \boldsymbol{Q}$, we get $\boldsymbol{T} = \boldsymbol{R}^H$. Thereby,

$$\text{SINR}_q = \frac{E_s P_q \boldsymbol{w}_q^H \boldsymbol{h}_q \boldsymbol{h}_q^H \boldsymbol{w}_q}{\sum_{i=1}^{q-1}\left[E_s P_i \boldsymbol{w}_i^H \boldsymbol{h}_q \boldsymbol{h}_q^H \boldsymbol{w}_i\right] + \sigma_q^2}. \tag{28}$$

V. Evalution and Performance Analysis

In this section, we evaluate the performance of ZFA, QR and that of MF. Afterward, we compare the proposed algorithms with the optimal one, the exhaustive search, via Monte-Carlos simulations. Hence we average the total system throughput over 100 random channel realizations. All base stations are assumed to transmit with identical power when using the exhaustive algorithm. The simulation parameters are taken from our previous work [9].

A. Total System Throughput versus SNR

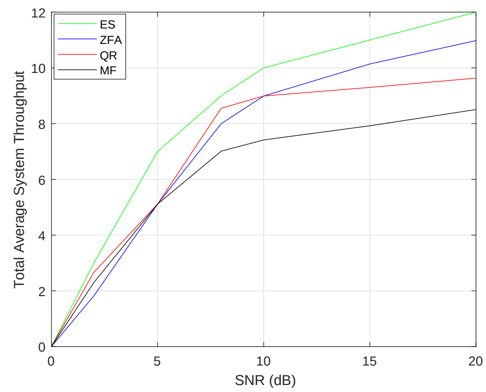

Figure 3. Total system throughput versus SNR in dB.

Figure 3 depicts the total system throughput for ZFA, QR and MF to be compared with the optimal solution ES. As shown in the low SNR regime from 0 to 5 dB, QR outperforms MF and ZFA due to high noise. Then, in middle SNR regime from 5 to 10 dB, we can detect that QR still outperforms but with remarkable progress for ZFA where the latter preceeds MF. For high SNR regime from 10 to 20 dB, it can be seen that QR still outperforms MF but ZFA preceeds the proposed QR solution. We could say that QR is a good solution for low and middle SNR range, while ZFA is more suitable for high SNR.

B. Average System Spectral Efficiency for Macro and Pico Cells versus SNR

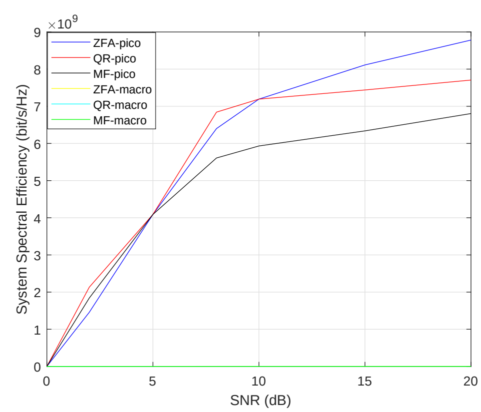

Figure 4. System spectral efficiency $bit/s/Hz$ for macro & pico cells.

Figure 4 we show the average system spectral efficiency in bit/s/Hz for both macro and picocells, aiming to clear out the importance of using mmWaves links in increasing the system throughput. TShe bandwidth used for macrocell is $B_w = 180kHz$ and that of picocells is $B_w = 800MHz$. As illustrated in Figure 4, the spectral efficiency of picocells by applying ZFA, QR and MF is roughly high x gigabit/s/Hz comparing to that of the macrocell x megabit/s/Hz. This due to the huge bandwidth proposed by mmWaves spectrum band. We can conclude that the concept of densification and heterogeneous network plays a role in maximizing the total system throughput and spectral efficiency. Because it gives the opportunity for users located at the border of the cell to be served and at the same time to reply to their requests.

C. System Execution Time

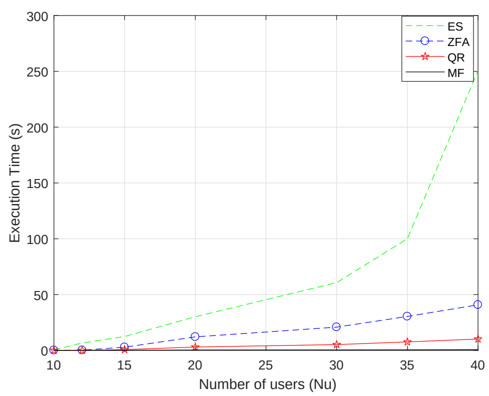

Figure 5. Execution time versus N_u users

Execution time is an important factor to be taken into consideration. Regarding Figure 5, execution time of QR does not exceed 5 s when $N_u = 40$ users, while ZFA needs four more time for the same number of users. But regarding the ES algorithm, its curve increases exponentially and tends toward infinity as the number of users increases. To interpret this result, we can see that QR outperforms ZFA and ES using the criterion of time, and it gives an acceptable throughput in low and middle SNR regime compared to optimal thus it could be considered as a trade-off algorithm between ZFA and ES.

VI. Conclusion

In this paper, we have formulated an optimization problem for a cross-layer design to optimize the hybrid beamforming. Notably, we have generated a new system model that supports the proposed scenario. The problem aims to maximize the total throughput for downlink multi-input-multi-output and orthogonal frequency-division multiple access system, for 5G heterogeneous networks (HetNETs). We have proposed different meta-heuristic algorithms to calculate the normalized beamforming vector, as we have applied the exhaustive search algorithm to get the optimal solution. Numerical results revealed that QR algorithm outperforms ZFA and MF in low and middle SNR regime while ZFA outperforms QR and MF as it provides higher system throughput. It is worth mentioning that QR outperforms ZFA, MF, and ES where it needs less execution time. As a conclusion, QR could be considered as a trade-off algorithm between ZFA and ES.

Acknowledgment

The project has been supported by the CNRS–GDR ISIS and doctor school MATHSTIC, to accomplish the collaboration between Vodafone chair and IRISA lab that tooks place at Dresden-Germany. Thanks for Abdul Karim Gizzini for all the technical support.

References

[1] M. Fallgren, B. Timus *et al.*, "Scenarios, requirements and kpis for 5g mobile and wireless system," *METIS deliverable D*, vol. 1, p. 1, 2013.

[2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[3] H.-L. Chiang, W. Rave, T. Kadur, and G. Fettweis, "Hybrid beamforming based on implicit channel state information for millimeter wave links," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 326–339, 2018.

[4] H.-L. Chiang, T. Kadur, and G. Fettweis, "Analyses of orthogonal and non-orthogonal steering vectors at millimeter wave systems," in *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2016 IEEE 17th International Symposium on A*. IEEE, 2016, pp. 1–6.

[5] S. Han, I. Chih-Lin, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5g," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 186–194, 2015.

[6] A. Zamani, S. Shojaee, O. Taghizadeh, and A. Schmeink, "Beamforming optimization with hybrid association in c-rans under a limited backhaul," in *Wireless Communication Systems (ISWCS), 2017 International Symposium on*. IEEE, 2017, pp. 240–245.

[7] G. Dimic and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Transactions on Signal processing*, vol. 53, no. 10, pp. 3857–3868, 2005.

[8] J. Zhang, A. Nimr, and M. Haardt, "Joint design of multi-tap filters and power control for fbmc/oqam based two-way decode-and-forward relaying systems in highly frequency selective channels." in *ICASSP*, 2015, pp. 2439–2443.

[9] A. Fadel, B. Cousin, and A. Khalil, "User selection in 5g heterogeneous networks based on millimeter-wave and beamforming," in *IEEE HPCC Conference*, 2018.

[10] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE journal on selected areas in communications*, vol. 32, no. 6, pp. 1164–1179, 2014.

[11] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE transactions on wireless communications*, vol. 14, no. 11, pp. 6481–6494, 2015.

[12] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave (mmwave) communications for 5g: opportunities and challenges. arxiv preprint," *Wireless Networks*, vol. 21, pp. 2657–2676, 2015.

[13] D. Yang, L.-L. Yang, and L. Hanzo, "Dft-based beamforming weight-vector codebook design for spatially correlated channels in the unitary precoding aided multiuser downlink," in *Communications (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.